Forecasting Output Growth Over the Medium-Term

Bernard Babineau and Nathanael Braun

Department of Finance Working Paper

2003-08

Department of Finance

Economic and Fiscal Policy Branch

104-03309/

Forecasting Output Growth Over the Medium-Term*

Bernard Babineau

Nathanael Braun†

Abstract

Medium-term (five-year) fiscal planning is complicated by uncertainty surrounding projections of real GDP. The slowdown in growth rate in the late 1970s and the recent upsurge in growth during the late 1990s provide lessons on how unanticipated movements in underlying trends can have a major impact on fiscal planning. This paper utilises linear and nonlinear (primarily) univariate models to examine the medium-term trends and uncertainty surrounding expected output growth. Specifically, we look at the "final" and "quasi-real" full sample forecasting performance of the models; that is, we use the most recently released, or "fully revised" real GDP data and consider forecasts using both fixed and fully recursive parameter values. In addition, we investigate the issue of whether (and how) the uncertainty around medium-term output growth has evolved since the mid-1960s.

Résumé

La planification fiscale à moyen terme est « compliquée » par l'incertitude entourant les prévisions du PIB. Le ralentissement économique à la fin des années 70 et la recrudescence de la croissance à la fin des années 90 démontrent l'impact important qu'un changement non anticipé de la composante tendancielle à sur la planification

* The authors would especially like to thank Doug Hostland and Jeremy Rudin for their helpful comments in the development of this paper. The views expressed in this paper are those of the authors and should not be attributed to the Department of Finance.

† Economic Studies and Policy Analysis Division (ESPAD), Department of Finance, L'Esplanade Laurier, 18th Floor, East Tower, 140 O'Connor Street, Ottawa, Ontario, K1A 0G5. Email: babineau.bernard@fin.gc.ca and braun.nathanel@fin.gc.ca.

fiscale. Dans cette étude, nous utilisons des modèles univariés linéaires et non linéaires pour estimer l'incertitude qui entoure nos prévisions du taux de croissance du PIB réel. Plus spécifiquement, on évalue la précision de nos modèles en temps « final » et « quasi-réel »; c'est-à-dire que l'on effectue des prévisions à partir de modèles avec des valeurs de paramètres fixes et pleinement récursifs en utilisant le plus récent échantillon du PIB. De plus, on regarde s'il y a eu changement dans le degré d'incertitude entourant la croissance du PIB depuis les années 60.

1 Introduction

The strong upsurge in the Canadian and American economies in the late 1990s has frequently been linked (especially in the US) to an increase in trend productivity growth associated with a bourgeoning "new economy," fuelled by the production and adoption of information and communication technologies (for example, see Jorgensen and Stiroh 2000 or Muir and Robidoux 2001). One of the difficulties for policy makers lies in assessing whether this change reflects a permanent or a transitory shift in the rate of productivity growth, and thus a corresponding change, *ceteris paribus*, in the rate of potential output growth. Although the type of shift may have little effect on forecasts of GDP growth during upcoming quarters it will have a progressively larger impact as the time horizon of the forecast increases, since if the recent increase proves to be permanent but forecasters deemed it temporary, forecasts of output growth would become progressively worse as the time horizon lengthens.[1] The objective of this paper is therefore to examine the magnitude of the uncertainty surrounding forecasts of medium-term output growth,[2] with the aim of assisting in the development of fiscal planning. In this paper we take the view that for forecasts made over the medium-term it is the permanent component of output growth that will dominate the estimates. Thus, in much of our discussion we will relate our uncertainty estimates back to recent discussion about the permanent (or potential) rate output growth.

[1] This is based on the assumption that the forecasts are unbiased initially.

[2] In this paper "medium-term" is defined as five years.

Why should fiscal authorities be concerned about the amount of uncertainty surrounding GDP growth, rather than just the actual forecasts? In the context of either long- or medium-term fiscal planning, uncertainty about whether the rate of output growth has changed permanently or temporarily can have significant implications for government's budgetary balances and planning. This uncertainty is relevant at shorter horizons, say one quarter to two years, but at shorter forecast horizons leading indicators and other variables may offer greater insight into the expected movements of GDP growth. Our use of a univariate forecasting framework is justified by the fact that these shorter run relationships are not as crucial in capturing the long-run dynamic of the GDP growth series. Since our concern lies in forecasting medium-term output growth, we focus more attention on how well phenomena like the 1970s slowdown and the recent late 1990s upsurge were captured by various forecasting techniques, rather than attempting to capture the cyclical movements of output growth.

Disentangling permanent and transitory shocks poses a difficult task, as does separating the permanent component of output growth from its cyclical component. Moreover, when the objective for policy makers is not only to assess the current permanent component of output growth, but also to determine the probable output growth of the economy over the next five years, misjudging current potential output may have undesirable consequences. Even if current potential output growth is estimated correctly, policy makers continue to face the difficulty of accurately predicting future changes in potential. According to one study by Muir and Robidoux (2001), recent changes in Canadian potential output growth have not been negligible; their estimates of potential output growth increase from 2.5 per cent in 1996 to 3.6 per cent in 2000 – over one percentage point in five years! Overestimation of output growth can have undesirable consequences for fiscal planning. It could inadvertently lead fiscal authorities to over commit to future government spending or tax cuts; thus, if realised output growth is lower than expected this would force policymakers to either backtrack on their previous commitments to program spending or else face a higher risk of a deficit outcome. The scope for reducing program spending, however, is constrained by the multi-year commitments. Thus, if program spending needs to be reduced to avoid a deficit outcome,

this may require the government to renege on previously promised program expenditures (see footnote 3).

In addition, the uncertainty around expected output growth has specific implications for setting an adequate level of "fiscal prudence" in the fiscal plan, a theme explored in recent research (Georges and Moreau 2001). Essentially, fiscal prudence refers to the money set aside to ensure against the risk of a deficit outcome. It has two components: a $3 billion contingency reserve and economic prudence. A budget balance of $3 billion or less at the end of the fiscal year is allocated paying down the debt, whereas any remaining surplus is either allocated to debt reduction and/or on additional program spending.

We can illustrate the potential impact of medium-term output growth uncertainty on fiscal planning in an example (table 1). In this example the fiscal plan is drawn up for the next five years at year zero; that is, before the government knows with certainty its revenue for any of the upcoming years. All program spending in any given year is thus made before any tax revenue is received. This implies that if realised government revenue turns out to be less than expected revenue, then the government must reduce its planned budget balance (in our example, this is set equal to $4 billion of fiscal prudence).[3] For example, in year one, if realised growth turns out to be less than what had been originally forecasted then the government is forced to reduce its year one budget balance. It then adjusts its spending in year two in order to attain the desired $4 billion budget balance for that year. The government makes this adjustment, however, assuming that growth will return to its originally forecasted level (thus only the level of program spending is adjusted, not its path). Thus (as is the case in table 1), if realised growth is below the originally forecasted growth rate, then the government will consistently miss its targeted budget balance and its level of program spending will have to be adjusted every year. Although this method of fiscal planning is slightly artificial it serves to illustrate the risks associated with medium-term uncertainty.[4]

---

[3] Debt reduction in our example is just the accumulation of the budget balance. Debt service payments are also taken as fixed, and are assumed to be known with certainty.

[4] See Georges and Moreau for more detail on the prudent fiscal planning process.

Table 1: Fiscal planning scenario without constraints on program spending
($ billion on a national accounts basis)

	Year 1	Year 2	Year 3	Year 4	Year 5
Planned (based on 3.5 per cent real output growth)					
Program spending	124.3	130.4	136.6	143.0	149.6
Budget balance	4.0	4.0	4.0	4.0	4.0
Debt reduction	4.0	8.0	12.0	16.0	20.0
Outcome (based on 3.0 per cent real output growth)					
Program spending	124.3	129.6	134.9	140.4	146.0
Budget balance	3.1	3.0	3.0	2.9	2.7
Debt reduction	3.1	6.2	9.1	12.0	14.7
Change in expenditures (outcome minus planned)					
Program spending	0.0	-0.8	-1.7	-2.6	-3.6
Budget balance	-0.9	-1.0	-1.0	-1.1	-1.3
Debt reduction	-0.9	-1.8	-2.9	-4.0	-5.3
Outcome (based on 2.0 per cent real output growth)					
Program spending	124.3	127.9	131.5	135.2	139.0
Budget balance	1.4	1.2	1.0	0.7	0.4
Debt reduction	1.4	2.7	3.6	4.3	4.7
Change in expenditures (outcome minus planned)					
Program spending	0.0	-2.5	-5.1	-7.8	-10.7
Budget balance	-2.6	-2.8	-3.0	-3.3	-3.6
Debt reduction	-2.6	-5.3	-8.4	-11.7	-15.3

Note: All calculations take into account that debt service payments vary with the amount of debt outstanding. Program spending and the budget balance are measured on a per year basis, but debt reduction payments are accumulated year-to-year.

Turning to table 1, if the standard deviation of forecasts of average real GDP growth over a five-year horizon is roughly 2 percentage points (this is later demonstrated to be a reasonable assumption), and average output growth is forecasted to be 3.5 per cent, this would imply that average real GDP growth over the upcoming five years has a 67 per cent probability of lying between 1.5 and 5.5 per cent (assuming a normal distribution). Looking at only the two cases where medium-term output growth is overestimated by 0.5 per cent and 1.5 per cent, a noticeable impact is found on the path of planned fiscal expenditures (see table 1).[5] When the average rate of output growth is overestimated by 0.5 of a percentage point the budget balance is consistently around one

[5] Note that in this example we assume that when the fiscal plan was made, real output growth was forecasted to be 3.5 per cent. The example we present here also assumes that program spending is flexible, and thus able to be reduced. Georges and Moreau (2001) examine alternative fiscal planning scenarios in which the amount by which program spending can be reduced is constrained by multi-period commitments.

billion dollars below what was originally anticipated, which roughly corresponds to the amount of economic prudence included in the fiscal plan. However, most of the $3 billion contingency reserve can still be allocated toward debt repayment. Over the entire five-year period, debt is reduced by $14.7 billion while program spending is reduced by a total of $9 billion relative to levels projected in the initial fiscal plan.

The fiscal situation deteriorates substantially when output growth is overestimated by 1.5 percentage points. Only $1.4 billion can be allocate to debt reduction in the first year; only $0.4 billion in the fifth year.[6] Debt is reduced by a total of $4.7 billion, while program spending is reduced by a total of $26 billion.

In assessing the uncertainty around medium-term output growth, this paper looks at the historical evolution of uncertainty surrounding forecasts of Canadian output growth over the medium-term (i.e., forecasting the average five-year growth rate). These forecasts are obtained using a several (predominantly) univariate models.

Section two begins by assuming fixed parameter values in order to present initial estimates of medium-term uncertainty and to illustrate the models and concepts used. This discussion is extended in section three, where more "accurate" estimates of uncertainty are obtained by allowing for uncertainty in the estimated models' parameters. In other words, in section two we assume that forecasters know the various models' parameter values, whereas in section three we relax this conjecture, and the parameters are estimated recursively. Section four discusses the evolution of medium-term output growth uncertainty over the sample period; and, section five concludes. All details regarding the models used are in annex one, while supplementary figures and tables appear in annex two.

2 A First Look at Forecasting Uncertainty

In this section we present the results of our forecasts of quarterly GDP growth. The models that we use to obtain medium-term output forecasts are all univariate models. It is possible that more general multivariate models may be desirable if one wishes to

[6] Given the original estimates of 3.5 per cent output growth , a deficit outcome would arise if realised output turned out to average only 1.8 per cent per year.

improve the models' forecasting performance, since such models would be able to incorporate leading and contemporaneous information. If the trend of output growth dominates medium-term forecasts then this additional information may have little to offer. However, if cyclical fluctuations drive a fair amount of the uncertainty over the medium-term, then multivariate methodologies may lessen this uncertainty. In this paper we begin under the assumption that the trend of output growth is likely what dominates average five-year output growth and thus, univariate estimation methods may adequately capture the dynamics of the real GDP series over the medium-term. At the very least, univariate methods provide a simple starting point for our analysis; they offer a benchmark upon which further research can build.

The approach taken in this paper is to look at the uncertainty in forecasting real GDP growth rates using univariate linear (autoregressive processes) and nonlinear models (breaking trend (BT), smooth transition autoregressive (STAR) and Hamilton). While used extensively in economics, linear models have been shown to be deficient in dealing with high volatility or the asymmetrical behaviour of some series, although this may be immaterial if the objective is medium-term forecasting. Nevertheless, most of the nonlinear models used in our paper aim to take into account the possible asymmetry in the GDP series. These nonlinear models stipulate that the process of the time series exhibits different dynamics according to the prevailing state of nature or regime (Hamilton 1989, Beaudry and Koop 1993, Potter 1995). This, in turn, implies that there is a different underlying "potential" growth rate depending upon the state of nature. Fiscal planning is thus obviously rendered more difficult since we have to assess not only the prevailing state of nature, but also the likelihood of a change in regimes in the medium term.[7] In addition, we include the linear model; since although it may not fit data as well as nonlinear models, we cannot rule out *a priori* that, in terms of forecasting, it does not perform as well as more complex models (see Stock and Watson 1998). The linear model also serves as a benchmark model, since it is encompassed by all of the other, more complex, models that we estimate.

All of the models used in this paper may be written as follows:

[7] Another implication of the nonlinear assumption is that the generating process driving the growth rate behaves differently to a negative shock than a positive one, which is not the case in the linear specification.

$$y_t = \sum_{i=1}^{r} x_t^T \beta_i F_i(S_t;\psi) + \varepsilon_t \quad \text{for } t = 1, \ldots, N \tag{1}$$

with $\sum_{i=1}^{r} F_i(S_t;\psi) = 1$; $F_i(S_t;\psi) \geq 0 \ \forall i$ and $\forall t$. Where β_i and ψ are $(p \times 1)$ and $(k \times 1)$ vectors of parameters respectively, r is the number of regimes, the $(p \times 1)$ vector of observed variables is x_t, and the $(q \times 1)$ vector of switching indicators is S_t. Finally, the function F_i indicates the state of nature or regime, and the error term, ε_t, is *i.i.d.* with mean zero and variance σ^2. Written in the form of equation (1), what distinguishes these models is the function F_i, and whether the state of nature is observable or not. Although the specific details regarding the estimation methodology used this paper are presented in annex one, we will briefly outline the intuition for the various models here.

Our linear model follows an autoregressive, AR(p), process, and in the framework of equation (1), it can be thought of as having only one regime. The BT model is an autoregressive model with an unknown number of break points, occurring at unknown dates. The intuition behind the BT model is that a change in the economy's "fundamentals" could lead to a shift in the mean of a series. Thus, we test for the presence of a discrete change in the trend of real GDP growth rate. Our test is twofold. First, we determine the most likely date where a break in trend may have occurred and then we test whether this break point is statistically significant. Our BT model assumes that both before and after the break point only the intercept of the AR process changes, and that the series follows the same order AR(p) process in both periods. In addition, the model requires a certain number of observations in order to test for a break point at either the beginning or the end of the sample. The latest date we can test for a break in Canada is 1997Q1 for the quarterly data.

Using the full data sample, we find one estimated break point in Canadian real GDP, occurring in 1976Q4. Trend real GDP growth was estimated to be 4.92 per cent before the break, and 2.80 per cent thereafter. Although this break point may appear later than the 1973 break date suggested by conventional wisdom, this is close to the 1976Q2 break found by Kichian (1999) using the Bai and Perron (1998) methodology. The slight difference in our estimated break point from that of Kichian can likely be accounted for by our use a different methodology to determine the break date, and our use of a longer,

more recent, data set. Although our BT model estimates a discrete break point in 1976Q4, there is little economic justification as to why a discrete change in the rate of output growth would have occurred in 1976; rather, this methodology just captures the existence of a slowdown around the mid- to late-1970s in a stylized manner. A more appropriate way to think of this break is to understand it as the (approximate) mid-point of a period of gradual slowdown in the rate of output growth. The dynamics of this change would thus be better captured by other estimation methodologies such as a state-space model, a STAR model, or an HP filter.

Turning to our estimated STAR model, instead of allowing time to determine the value of the function F_t (or the state of nature) as with the BT model, in the STAR model F_t is the value of some observable variable.[8] The STAR model can be thought of as essentially a generalization of a threshold model. A simple example will illustrate the intuition behind these models. Suppose that a negative growth rate of GDP the last quarter offers information on the current state of the economy. A pure threshold model would indicate that the economy is in a high growth state if we observe a positive growth rate the previous quarter and in a low growth regime otherwise. The STAR model would allow for transitional regimes between the high and low growth extreme regimes. Instead of just simply adding an intermediate regime as in most threshold models, the STAR model assumes a continuum of intermediary regimes. We use a nonlinearity test to suggest the most likely lag value for the switching variable S_t. In this paper we considered lagged values of Canadian GDP growth, the change in the bank rate, and US GDP growth. Given this test we select the fifth lag of the difference of the bank rate, as well as the one-quarter lag of US GDP growth rate.

Hamilton (1989) proposed a regime-switching model based on an unobservable state of nature S_t. This is again similar to the idea of a threshold model except that in this case the current regime is no longer determined by an observable variable such as lagged GDP, but instead it is based on an unobservable state of nature.[9] Though the states are

---

[8] Thus, our STAR model is not strictly univariate.

[9] For the full sample, it is assumed that S_t follows a first-order Markov process. The first-order Markov assumption enables us to solve the maximisation of the joint likelihood function of the economic series and the unobserved states of nature.

unobservable, we can nonetheless infer the probability of being in one state by looking at previous observations.

To assess the fit of the various nonlinear models, Terasvirta and Anderson (1992) suggest that the STAR specification should be rejected if the ratio of the residual variance of the STAR model over the linear model is greater than 0.9, or if the reduction in the standard deviation obtained by going from the linear time-series model to a STAR model is less than 10 per cent. In general, we will use this simple rule of thumb to determine the acceptance of any nonlinear model over the simpler linear specification.[10]

In forecasting Canadian real GDP growth the forecast of the series, y, is simply the conditional mean: $\hat{y}_{t+h} = E(y_{t+h} \mid I_t)$, where h refers to the horizon, or the number of periods forecasted ahead.[11] Our estimate of "uncertainty" is the root mean squared error (RMSE), which measures the deviation of the forecasted values ($\hat{y}$) from the actual values (y).[12]

In addition to forecasting using the above models we also perform *ad hoc* "forecasts" using the Hodrick-Prescott (HP) filter (Hodrick and Prescott 1997).[13] The principal reason for doing this is the widespread use of the HP filter for trend-cycle decompositions. The estimated HP potential GDP growth is sometimes treated (explicitly or implicitly) as equivalent to the rate of expected GDP growth. Thus, when we forecast the HP filter we are simply projecting the estimated trend at time t out over

---

[10] Using this variance ratio rule, the ratio for the various models is as follows: BT, 0.905; STAR (bank), 0.936; STAR (US), 0.897; Hamilton, 0.572. We retain the STAR (bank) model even though it is above our "rule of thumb" cut-off for the "final" STAR estimates, since (on average) it is not below the cut-off for our quasi-real estimates (see figure A3).

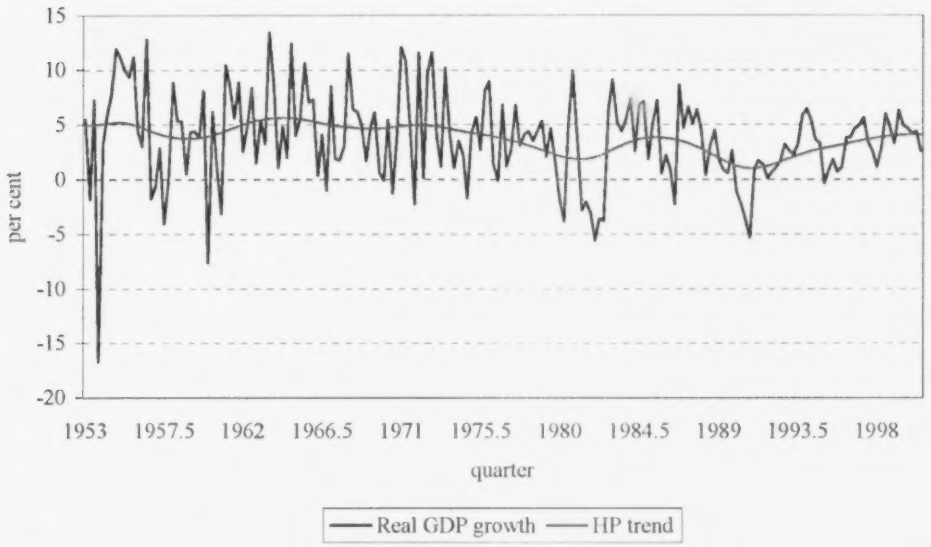
[11] Note that $E(F(z_{t+h}) \mid I_t) \neq F(E(z_{t+h} \mid I_t))$, and while the former is the optimal forecast in the STAR model, we use the latter as our conditional forecast. In addition, note that according to our specification, the STAR models behave as a linear model for forecasting purposes with parameter values determined by last available value of the switching variable in the information set.

[12] The RMSE is defined as $RMSE = \sqrt{\frac{1}{T}\sum_{t=1}^{T}(\hat{y}_t - y_t)^2}$.

[13] Since the HP filter is a two-sided filter it has well-established problems with estimated trends at the end of a sample period. In estimating the HP trend of our series the HP filter becomes a one-sided filter toward the beginning and the end of the sample period. In addition, recall that the HP filter decomposes a series, y, into an additive cyclical (y^C) and permanent (y^P), where

$$\left\{y_t^P\right\}_{t=0}^{T} = arg\ min \sum_{t=1}^{T}\left\{\left(y_t - y_t^P\right)^2 + \lambda\left[\left(y_{t+1}^P - y_t^P\right) - \left(y_t^P - y_{t-1}^P\right)\right]^2\right\}$$

Figure 1: Real GDP growth



the subsequent five years.[14] We are not suggesting that this method of estimating the HP trend or producing HP forecasts is optimal or even desirable; however, despite its *ad hoc* nature it is interesting to note (in upcoming sections) how well it performs. Like the Linear model, it also serves as a benchmark against which we can compare our other models.

Finally, before discussing the results of our estimation, note that the data used in this paper are quarterly Canadian real GDP from Statistics Canada over the period of 1953Q1 to 2000Q4. These data are presented in figure 1, which includes the HP trend in order to aid the reader in seeing the underlying evolution of the series. Since the new Fisher Chain weighted data are available only back to 1981, we have spliced the chained series together with a historical GDP series in order to obtain a longer dataset. Although this introduces a break in our dataset, we do not feel that this has had a significant

[14] Note that in our estimation of the HP filter, we set lambda equal to 1600. It should be noted that we do not suggest that this interpretation of the HP trend (as equivalent to the expected growth rate) necessarily has any economically or statistically justification.

Table 2: Final medium-term (five-year) RMSEs

Final	Linear	STAR (bank)	STAR (US)	BT	Hamilton	HP Filter
1960-1995	1.341	1.374	1.168	0.974	1.318	1.241
1965-1995	1.399	1.416	1.194	0.939	1.370	1.267
1970-1995	1.350	1.350	1.237	0.976	1.284	1.341

distortionary impact on our results.[15] In looking at the data, note that Canadian real GDP growth was more volatile in every decade prior to 1990 than it was during that decade (see figure 1). In addition, the HP filter suggests a gradual reduction in the mean of real GDP growth during much of the 1970s (this is also picked up our switching models). Major recessions were also experienced in both the early 1980s and the early 1990s, both of which have a noticeable impact of the HP trend estimates.

In order to produce only one forecast per year, we only utilise our models forecasts' that are made using fourth quarter data (this allows us to have only one forecast made per year, and also assists in minimising the computational burden of several of our models). Our various models then forecast quarterly real GDP growth over upcoming twenty quarters (or five years), however, since the concern of policy makers (at least for fiscal planning) is not so much the growth rate twenty quarters ahead, but rather the average growth rate over the upcoming twenty quarters (five years), it is this latter definition on which this paper will focus.[16] Thus, we average twenty quarters of our quarterly forecasts and compare each of these five-year forecasts to the average actual real GDP growth.

The in-sample forecasting performance of our five models and the HP filter are presented in table 2 for different sample periods. These estimates are called "final" since they use the parameters estimated from the "final" full sample dataset.[17] Our data begin

[15] There is little difference in the average five-year growth rates between the previous Laspayres GDP data and the Fisher data that we use. The volatility of the series also hardly changes. Between 1981Q2 and 2000Q4 the standard deviation of the (quarterly) Laspayres series is 3.205, while it is 3.245 for the Fisher data that we use.

[16] Thus, one-year ahead forecasts are the average of the forecasts for the next four quarters, while the two-year ahead forecasts are the average of the forecasts for the next eight quarters, *et cetera*.

[17] This terminology derives from the real-time literature (for example, Orphanides and van Norden 1999). In real-time analysis the term "final" makes more sense in that case since there are multiple datasets, and

in 1953; however, we present the RMSEs beginning in 1960, 1965 and 1970, because some nonlinear models require a certain amount of data to "adequately" estimate their parameters (although this does not impact the final RMSEs the importance of this will become evident in the next section).

Using final parameter values, the BT model is found to have the best medium-run forecasting performance, followed by the STAR (US). The next-best forecaster is the Hamilton model, although, surprisingly the Linear, the STAR (bank) model and the *ad hoc* HP filter perform similarly.[18] Thus, overall we can say that the range of uncertainty varies from approximately 0.9 to 1.4 percentage points. This implies that if our mean forecast of real GDP growth was 3.5 per cent for the upcoming five years, we would have a probability of 67 per cent of being between 2.6 and 4.4 per cent (using the RMSE of 0.9) or between 2.1 and 4.9 per cent (with a RMSE of 1.4), assuming a normal distribution. These models produce a wide range of forecasts (figure 5 below). For example, in forecasting real GDP growth over the 2001 to 2005 period forecasts range from 2.8 per cent (from the BT model) to 4.1 per cent (from the HP filter), with the remaining forecasts in the range of 3.1 to 3.8 per cent. Note that the mean medium-term forecast of 3.5 per cent is consistent with the estimate of potential output growth obtained by Muir and Robidoux (2001).

3 Quasi-Real Forecasting Uncertainty

In this section we extend the results of the previous section to allow for the case where the parameter values are estimated recursively. That is, whereas in the previous section

---

"final" refers to the most recently available, or "fully revised" one. Despite having only one dataset we will also adopt this terminology.

[18] Estimates of the uncertainty (or forecasting performance) produced by the various models over the one to five year horizons are found in table A1. The forecasting performance (ranking) of the various models also changes according to the forecasting horizon. For example, the BT model performs quite well at longer horizons, but does comparatively less well over shorter forecasting horizons.

In addition, note that the RMSE for all models decreases as the forecasting horizon increases. This may seem counter-intuitive at first, since one would typically expect uncertainty to increase at longer horizons, especially when a series is subject to both permanent and transitory shocks. However, it should be recalled the magnitude of the RMSE is contingent on the standard deviation of the series, which is being forecasted. Since we are forecasting average real GDP growth rates over the one to five year horizons, we are not strictly speaking forecasting the same series. Note that although the mean of all five series – the one, two, three, four, and five year average growth rates – are the same, their respective standard deviations are decreasing.

Figure 2: Hodrick-Prescott trend GDP growth

we treated the parameter values as known, whereas in this section we assume that the policy maker does not know the "true" parameter values. Thus, in estimating parameters recursively, this section approximates more closely the actual uncertainty being faced, in the sense that we use the parameter value at time t that would have obtained using data up to t. This, however, does not allow us to completely capture the uncertainty facing the policy makers since we are not using a real-time data set – i.e., a data set with the specific data series that were available at each point in time.

To demonstrate the importance of using quasi-real estimates forecasts, we will begin this section by presenting a simple example using the HP filter. Figure 2 compares the final and quasi-real HP estimates of trend output growth (keep in mind that these are the same as the HP forecasts in this paper, since we use trend output growth to forecast medium-term output growth).[19] For example, this graph tells us that in 1990 we would

[19] The quasi-real trend value at time t corresponds to the HP potential output at time t based on the information set at time t. The final HP trend series is the potential output at t produced using the full sample (T observations). Thus, essentially the final estimates of the trend at t are made with foresight (they know what future output growth looks like), whereas the quasi-real estimates only have information up to date t. These estimates, however, are not really quasi-real in the sense of parameter uncertainty since we are using the same lambda (1600) for all estimates. The difference in the HP quasi-real and final estimates is due to the way that the filter treats the end-of-sample estimates. Recall that since the filter is essentially a two-sided moving average filter, it switches to being one-sided as it approaches the end of the sample period.

Table 3: Quasi-Real Medium-term (five-year) RMSEs

	Linear	STAR (bank)	STAR (US)	BT	Hamilton	HP Filter
Quasi-Real						
1960-1995	2.009	2.375	2.074	2.005	2.258	1.820
1965-1995	1.646	2.003	1.531	1.666	1.846	1.835
1970-1995	1.748	1.974	1.412	1.743	1.645	1.951

have thought that permanent output growth was 4 per cent; whereas when the HP filter is applied to the data available in 2000, the permanent component in 1990 is closer to 2 per cent. This raises the question of how much the 4 per cent estimate of trend output growth produced by the HP filter for 2000 could change in upcoming years.

Likewise, we carry out a similar exercise with our various models in this section, re-estimating the parameter values every year and then using these parameter values to forecast five years ahead (table 3).[20, 21] Comparing this quasi-real case with the more "artificial" final forecasts, notice that all models' forecasting performance deteriorate markedly. These estimates are a better reflection of the uncertainty faced than those produced in table 2, since at least part of the uncertainty faced in producing medium-term forecasts also lies in estimating the correct parameter values. Whereas the BT model was the best performer when the parameter values were estimated over the full sample, it is now the STAR (US) model that performs the best when the parameters are estimated recursively. The BT model continues to perform comparatively well, and the Linear model also does surprisingly well.

[20] Once again, estimates of the uncertainty produced by the various models over the one to five year horizons are found in table A1.

[21] Given that our criterion for model selection is the ratio of the residual variance, and that this ratio is derived from the fit of a model, we plot the quasi-real standard errors and the ratios of the residual variances in figures A1 and A2 respectively. It is interesting to note that the fit of all models (with the exception of the Hamilton model) more or less continuously improve in their fit over the sample period. In addition, their "improvements" follow a quite similar pattern.

Looking at figure A4, the Hamilton model is always the best performing model (except in 1982), in terms of standard error. Likewise, we would always accept the Hamilton and STAR (level) models in favour of the Linear model. The STAR (difference) model would have more or less been accepted up until the 1990s, and the BT model would never have been accepted, except maybe during the 1990s when it hovered around the rule-of-thumb cut-off point of 0.9.

Notice that as we shorten our sample period over which the RMSE is calculated the Hamilton model and both STAR models do progressively better. This early poor forecasting performance may be due to the sensitively of these regime-switching models to small data samples; that is, given a small number of observations (and therefore the small likelihood of observing the regime shift) these models may be more sensitive in estimating parameter values. One of the difficulties in using these regime-switching models is finding a global maximum of the likelihood function. This may be less a problem in the case of the STAR model since in each sub-period we can perform a grid-search in order to get "appropriate" starting values for the parameters in the iterative optimization procedure. This is more problematic in the case of the Hamilton model since the grid-search would have to be carried over a large number of parameters.

Finally, although we take parameter uncertainty into account in this section, our analysis does not account for data uncertainty. That is, the real GDP data set that we use in this paper does not contain the same estimates of real GDP growth that were initially available to policy makers for any given year (or quarter). Statistical agencies tend to make some revisions to estimates of real GDP for several quarters after their initial release, and further revisions are made on a less frequent basis, such as when real GDP is rebased, or when particular components are redefined statistically. Accounting for this data uncertainty can be done by using real-time data sets (i.e., a data set with the specific data series that were available at each point in time). In contexts other than forecasting uncertainty, recent US research has found that the use of real-time data can have a noteworthy impact on the evaluation of past experiences and performance.[22]

4 Evolution of Forecasting Uncertainty

In this section, we extend the results reported in the two previous sections to look at the evolution of both final and quasi-real uncertainty over time. In doing so, we can observe whether incertitude has remained relatively consistent, or whether there have been large variations over time.

---

[22] Stark and Croushore (2001) present the main US real-time data set and illustrate the impact that real-time data can make on the forecasts yielded by various models. Orphanides and van Norden (1999) look at the impact that real-time data can have when attempting to evaluate the current output gap. Their paper uses real-time data to evaluate, *ex post*, the ability of various models to accurately estimate the output gap.

Figure 3: Final rolling RMSE

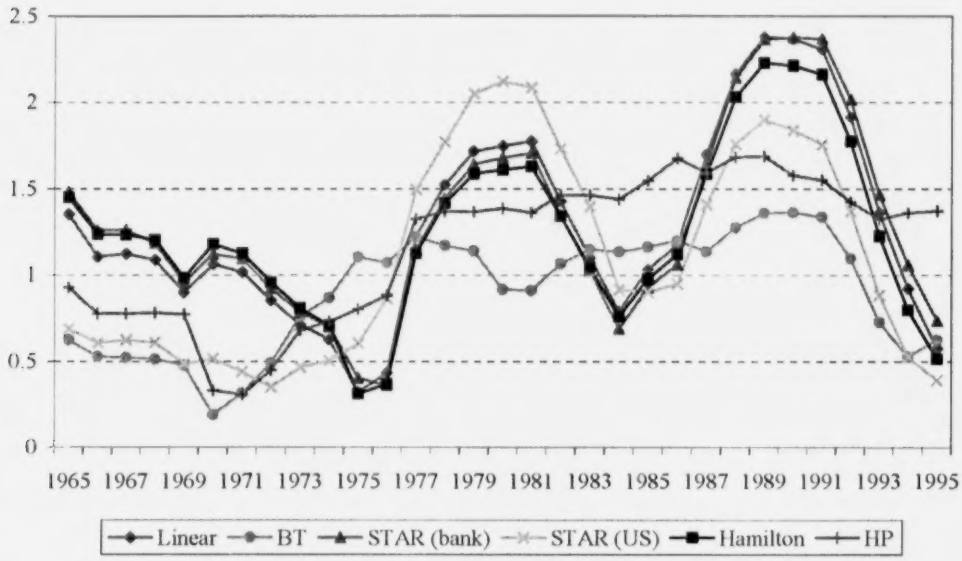
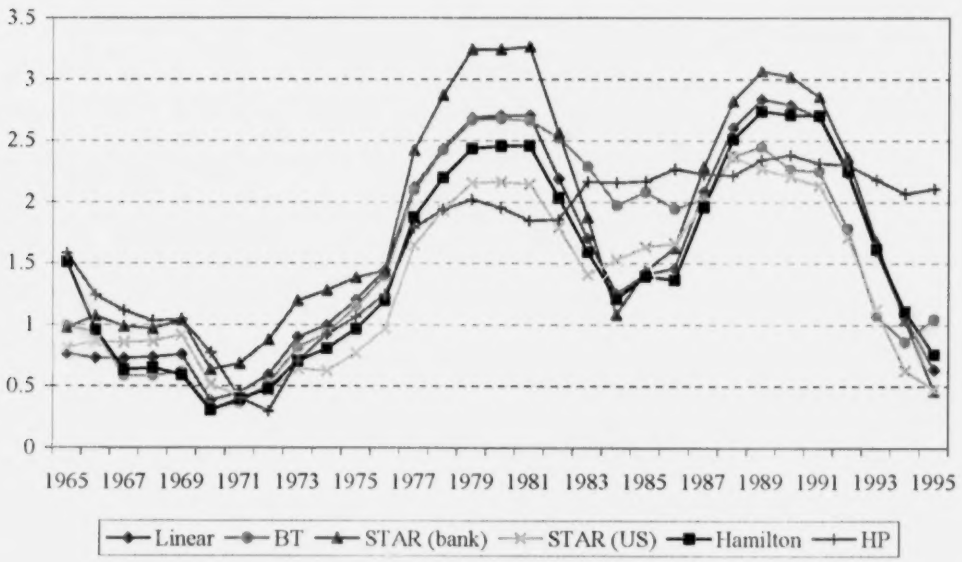


Figure 4: Quasi-real rolling RMSE



Figures 3 and 4 present rolling RMSEs for our final and quasi-real forecasts. By "rolling" we mean the RMSE calculated over the previous five years – i.e., the rolling RMSE in 1995 is equal to the RMSE calculated over the period 1991 to 1995.

Turning attention first toward figure 3, notice that the RMSEs of the Hamilton, Linear and STAR (bank) models are very highly correlated (cross-correlations are presented in tables A2 and A3). These RMSEs decrease throughout most of the 1960s and 1970s, while the other models' forecasting errors remain relatively low, or also decrease slightly over this same period. The performance of these three models deteriorates rapidly in the late 1970s, improves briefly during the early to mid 1980s; they deteriorates once again with the early 1990s slowdown, but they all steadily improve since 1990. In the early 1970s the forecasting errors of the HP filter and the BT model (and to a lesser extent the STAR (US) model) steadily increase until around 1980; the BT model's RMSEs decline again in the 1990s.

Turning to the quasi-real rolling RMSEs (figure 4) notice that although the Linear, STAR (bank) and Hamilton models all still follow a similar pattern, they are not as quite highly correlated (table A3). Specifically, the STAR (US) model and the BT model's RMSEs are now more highly correlated with the movements of the Linear, STAR (bank) and Hamilton models than they were for the final forecasts. Another noteworthy difference is that while the final RMSEs suggest that uncertainty decreases over the 1960s and much of the 1970s, the quasi-real RMSEs decrease slightly over the late 1960s, but they then all increase steadily from the early 1970s to the early 1980s. This finding fits with the observed increase in uncertainty over this period. Over the course of the 1970s and early 1980s the uncertainty around expected output growth increased, driven, at least in part, by the two oil crises and the productivity slowdown that occurred over this period. Similarly, quasi-real uncertainty peaked again around the major slowdown in the early 1990s, but it has noticeably improved since that time. Thus, although one may expect the trend of output growth to dominant in medium-term forecasts, it appears that a large amount of our estimated forecasting uncertainty is being driven by our various models inability to pick up the deep recessions of the early 1980s and early 1990s.

Figures 5 and 6 present the medium-term final and quasi-real forecasts produced by our different forecasting methodologies. Several features are evident from these forecasts, however, overall it is remarkable to note how similar are most of the forecasts produced by the different models. The discrete break in the full sample BT model

Figure 5: Final medium-term forecasts

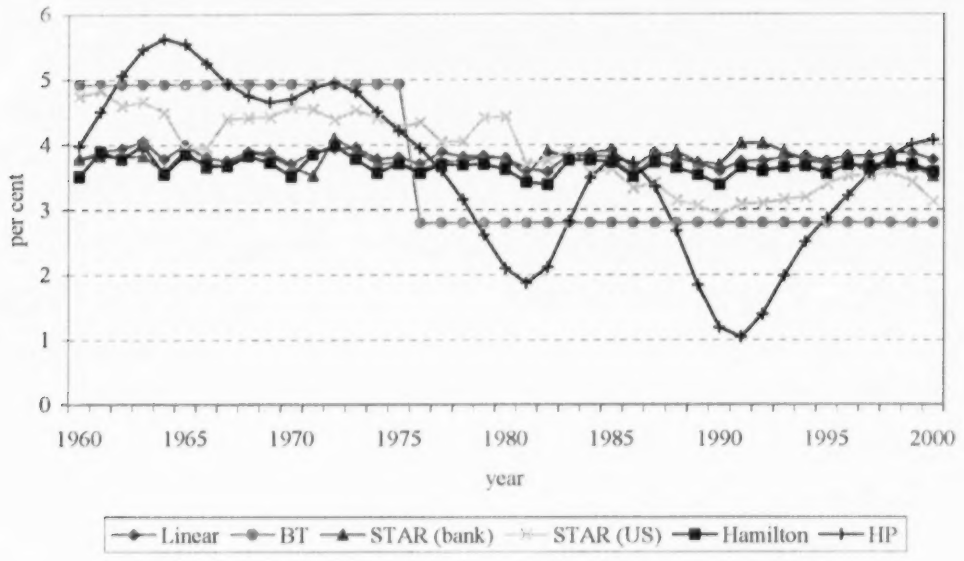
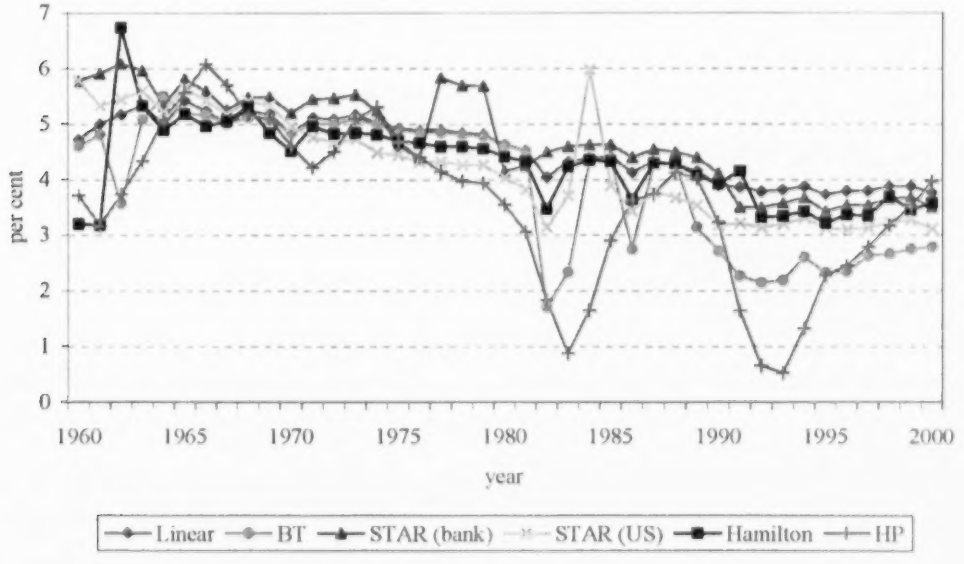


Figure 6: Quasi-real medium-term forecasts



produces better forecasts than the other models. Although when we move to the quasi-real forecasts, which the incorporate uncertainty regarding the BT's break point, this results in BT forecasts become quite similar to those produced by the other models (at

least up until the 1990s). Thus, even though the BT model's final forecasts may appear quite different from those of the other models, the forecasts that the BT model would *actually* produce (i.e., quasi-real forecasts) are not qualitatively much different from those of other models.

Another possible undesirable aspect of using the BT methodology in practice is the difficulty it has in detecting a break in the recent past. For example, despite the fact the slowdown in Canadian output growth occurred around the mid-1970s, a break was only first detected in 1982Q1, and it was only consistently found after 1989Q2.[23] Likewise, if trend output growth has actually increased in the late 1990s it is unlikely to be detected using the BT methodology for at least several years to come. In part, decreasing the desired critical value for detecting a break point could mitigate this issue, but it will not eliminate this problem. Next to the BT model, the STAR (US) model best forecasts output growth during the 1970s slowdown.

Although the forecasting performance of the HP filter is surprisingly good, when one looks at the forecasts that it produces it becomes evident that the HP filter (at least in this form) would be an undesirable methodology for medium-term forecasts. Since the forecasts that it produces are more subject to large cyclical swings than other models', they are unlikely to allow for consistent fiscal planning. However, even though the HP forecasts perform poorly during "good times" (the mid to late 1980s and 1990s), they do better than the other methodologies during recessionary periods (although some of this better performance disappears with the quasi-real estimates).

In summary, quasi-real forecasting performance varies slightly across the different models; the main source of the uncertainty estimated in tables 2 and 3, however, is the large variations in RMSEs of all models at different points in time. To a large extent, most of our uncertainty surrounding medium-term forecasts resides in our models' inability to detect the deep recessions of the early 1980s and 1990s (with the exception of the HP filter).

[23] In addition, the estimated break date has not consistently been 1976Q4 since 1989Q2.

5 Conclusion

The main objective of this paper is the gauge the medium-term uncertainty surrounding expected output growth. The RMSEs obtained from linear and nonlinear models forecasts range from around 1.4 to 2.3 percentage points when we allow for parameter uncertainty (the quasi-real forecasts). Much of the uncertainty coincides with the two major Canadian recessions of the early 1980s and 1990s. Toward the end of our sample period, however, forecasting uncertainty appears to have decreased to levels last experienced in the late 1960s and early 1970s, although this, in part, may be due to our small sample size (five forecasts for the rolling RMSEs). As demonstrated in the example in the introduction, however, even a small amount of uncertainty can have significant implications for fiscal planning.[24] Hence, this paper demonstrates the magnitude of uncertainty surrounding expected medium-term real output growth, and it re-emphasises the important of incorporating flexibility, such as is discussed in Georges and Moreau (2001), into fiscal planning in order to take account for this degree of incertitude.

Several noteworthy findings emerge from this paper. First, with regard to fiscal planning, there appears to be a fair degree of uncertainty around medium-term output growth. Thus, the overestimation of output growth by one per cent (as in the example in the introduction) is not unreasonable. Accounting for the presence of uncertainty plays an important role for policy makers. This uncertainty is especially important for medium-term fiscal policy, where the time horizon of concern is over the next five years. The appropriate estimation and inclusion of uncertainty in the development of medium-term fiscal policy will thus assist in ensuring sound fiscal planning through the adequate allocation of fiscal prudence.

Second, the inclusion of parameter uncertainty (our quasi-real distinction) is quite important for measuring uncertainty. Not only do uncertainty estimates increase, but the relative forecasting performance of the various models also changes considerably over time. Models that forecast quite well when their parameter values are assumed to be known deteriorate significantly when their parameters are recursively estimated.

---

[24] Given our range of RMSEs, and assuming that they are normally distributed, there is approximately a 15 per cent probability that average five-year output growth could be overestimated by 1.5 percentage points (as was the case in the example in the introduction).

Third, given that particular models appear to produce better forecasts, it is quite important to look at the evolution of the various models' forecasting performance to determine whether it is particular periods of performance that are driving the RMSEs. Our anecdotal evidence that while many models produced RMSEs between 2 and 3 during the late 1980s and early 1990s, these same models were producing RMSEs between 0.5 and 1.5 prior to the mid-1970s and after 1993. Given these large fluctuations over time, what are the implications for prudent fiscal planning? Are their any, more structured ways that can be used to estimate how uncertainty changes, or evolves? Our paper assumes that the uncertainty surrounding forecasts of output growth is symmetric. Although this is a common assumption, it may or may not be realistic. Hence, it may beneficial for future research to investigate the distribution of uncertainty around forecasted output growth (and how, or whether, this distribution changes over time).

Finally, there are several other ways in which this paper's results could be further explored. It is possible that multivariate models would yield more accurate forecasts of medium-term output growth, and thereby reduce uncertainty. This is alluded to by strong performance of the STAR (US) model in this paper. The inclusion of additional (or leading) information may improve our ability to forecast medium-term output growth, and thereby decrease uncertainty.

In addition, this paper may underestimate uncertainty, since real-time output growth data are not being used. Using real-time data would allow us to take into account not only parameter uncertainty but also the uncertainty surrounding future data revisions. It is possible that this last source of uncertainty is not negligible in magnitude (Stark and Croushore 2001). Thus, although our quasi-real forecasts replicate real-time forecasts in that they are fully recursive, they do not take into account revisions to the data by statistical agencies, a frequent and statistically important occurrence. Future work on forecasting performance, forecasting uncertainty, and policy rules could benefit by taking into account data uncertainty, thereby improving efficacy of empirical findings to policy makers (although at present there is still a paucity of real-time data sets for Canada).

Bibliography

[1] Andrews, Donald W.K. 1993. "Tests for Parameter Instability and Structural Change with Unknown Change Point," *Econometrica*, 61, pp. 821-856.

[2] Andrews, Donald W. K. and Werner Ploberger. 1994. "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62(6), pp.1383-1414.

[3] Beaudry, Paul and Gary Koop. 1993. "Do Recessions Permanently Change Output," *Journal of Monetary Economic*, June, pp.149-163.

[4] Bai, Jushan and Pierre Perron. 1998. "Estimating and Testing Linear Models with Multiple Structural Changes," *Econometrica*, 66(1), pp.47-78.

[5] Georges, Patrick and Nicolas Moreau. 2001. "Prudence, Scope for Discretionary Measures and Conditional Rules in Medium-Term Fiscal Planning," *draft*, Economic Studies and Policy Analysis Division, Department of Finance.

[6] Granger, Clive W. J., and Timo Terasvirta. 1993. *Modelling Nonlinear Economic Relationships*. Toronto: Oxford University Press.

[7] Hamilton, James D. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica*, 57(2), pp. 357-384.

[8] Hansen, Bruce E. 2000. "Testing for Structural Change in Conditional Models," *Journal of Econometrics*, 97(1), pp.93-115.

[9] Hodrick, Robert J. and Edward C. Prescott. 1997. "Postwar US Business Cycles: An Empirical Investigation," *Journal of Money, Credit, and Banking*, 29(1), pp.1-16.

[10] Jorgenson, Dale W. and Kevin J. Stiroh. 2000. "Raising the Speed Limit: US Economic Growth in the Information Age," *Brookings Papers on Economic Activity*, 1, pp.125-235.

[11] Kichian, Maral. 1999. "Measuring Potential Output within a State-Space Framework," Bank of Canada *Working Paper* 99-9.

[12] Muir, Dirk and Benoît Robidoux. 2001. "Information Technology and the US Productivity Revival: Is Canada Lagging Behind?," Department of Finance Working Paper (forthcoming).

[13] Orphanides, Athanasios and Simon van Norden. 1999. "The Reliability of Output Gap Estimates in Real-Time," Board of Governors of the Federal Reserve System (US) *Finance and Economic Discussion Series* No. 1999-38.

[14] Potter, Simon. 1995. "A Nonlinear Approach to US GNP," *Journal of Applied Econometrics*, May, 10, 109-125.

[15] Stark, Tom and Dean Croushore. 2001. "Forecasting with a Real-Time Data Set for Macroeconomists," *Federal Reserve Bank of Philadelphia Working Paper* No. 01-10.

[16] Stock, James H. and Mark W. Watson. 1998. "A Comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series," *NBER Working Paper* No. 6607.

[17] Terasvirta, Timo. 1994. "Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models," *Journal of the American Statistical Association* 89(425), pp.208-218.

[18] Terasvirta, Timo. 1995. "Modelling Nonlinearity in US Gross National Product 1889-1987," *Empirical Economics*, 20(4), pp.577-97.

[19] Terasvirta, Timo, and Heather M. Anderson. 1992. "Characterising Nonlinearities in Business Cycles Using Smooth Transition Autoregressive Models," *Journal of Applied Econometrics*, 7(0) Suppl., pp.S119-S136.

Annex 1: Estimation Methodologies

This annex briefly presents the four types of univariate regime-switching models (plus the linear model) used to derive medium-term forecasts of output growth. These are all models for which the generating process of a series is function of some state of nature. The models used in this paper are a breaking trend (BT) model, a smooth transition autoregressive (STAR) model, and a Hamilton regime-switching model. All of these models may be written in the form:

$$y_t = \sum_{i=1}^{r} x_t^T \beta_i F_i(S_t;\psi) + \varepsilon_t \quad \text{for } t = 1,...,N \tag{A1}$$

with $\sum_{i=1}^{r} F_i(S_t;\psi) = 1$; $F_i(S_t;\psi) \geq 0 \ \forall i$ and $\forall t$. Where β_i and ψ are respectively $(p \times 1)$ and $(k \times 1)$ vectors of parameters and r is the number of regimes; the $(p \times 1)$ vector of "observed" variables is x_t, and the $(q \times 1)$ vector of switching indicators is S_t. Finally, the function F_i indicates the state of nature or regime, and the error term, ε_t, is *i.i.d.* with mean zero and variance σ^2. Given that all of the aforementioned switching models can be written as (A1), what will distinguish these models is essentially the function F_i, and whether the state of nature is observable or not. In addition to presented to models, we also present the estimated "final" parameters values (see section 2).

Linear Model. Our estimated final linear model follows a second-order autoregressive process (AR(2)), as determined using a number of criteria, such as the Schwartz and Akaike criteria, and the t-statistics of the last lagged value. The mean growth rate implied by this model is 3.81. The estimated model is:

$$y_t = 2.557 + 0.235 y_{t-1} + 0.094 y_{t-2} + \hat{\varepsilon}_t, \text{ with } \hat{\sigma} = 1.955.$$

Breaking Trend Model. In the BT model, the indicator function F_i is a function of time, and takes on two discrete values {0,1}. The simplest case is to assume two regimes, which enables us to express F_2 in terms of F_1 (i.e., $F_2 = 1\text{-}F_1$). We would then have the following model:

$$y_t = x_t^T \beta_2 + x_t^T (\beta_1 - \beta_2) F_1(t;\psi) + \varepsilon_t \text{ with } F_1(t;\psi) = \begin{cases} 1 & \text{if } t \le T_\psi \\ 0 & \text{otherwise} \end{cases} \quad \text{(A2)}$$

The estimation of the BT model would be trivial if the date of the break (T_ψ) was known, since it would simply require the addition of a dummy variable to the linear time series model. Bai and Perron (1998) put forth an extensive discussion on the estimation and testing of the breaking trend model when the break dates are unknown.

We calculate three test statistics – AveF, SupF, and ExpF – to determine the significance of the break; if there are discrepancies we rely on ExpF as the final criterion for break-significance. SupF is from Andrews (1993), while ExpF and AveF are from Andrews and Ploberger (1994). For a discussion of these tests see Hansen (2000). The estimated BT model is:

$$y_t = 2.149 + 1.635 I_t + 0.216\, y_{t-1} + 0.023\, y_{t-2} + 0.130\, y_{t-3} - 0.137\, y_{t-4} + \hat{\varepsilon}_t,$$
$$\text{with } \hat{\sigma} = 1.907.$$

Smooth Transition Autoregressive Model. An alternate approach to the BT model is to use a continuous function $F_t: \Re \rightarrow [0,1]$ called the transition function. These types of models are called STAR models, and have largely been advocated by Granger and Terasvirta (1993) and Terasvirta (1994; 1995). While in the BT model there are r possible regimes or breaks, STAR models up until now have been expressed solely in terms of two "extreme" regimes with continuous fluctuations between these two extreme, or limit, regimes.

The functional form for the transition function, F, used in this paper is the logistic function:

$$F(S_t; \gamma, c) = \frac{1}{1 + exp(-\gamma(S_t - c))} \quad (\gamma > 0). \quad \text{(A3)}$$

The logistic specification defines one regime to exist when the switching variable is above a critical threshold value, and the other to exist when this variable is below the cut-off point. A maximum likelihood approach is used to estimate the STAR model.

The strategy is to concentrate out the likelihood function, permitting a grid-search approach to narrow the starting values of certain parameters.

Allow

$$z_t(\psi)^T = [\, x_t^T \times F_1(S_t; \psi) \ldots \; x_t^T \times F_r(S_t; \psi) \,],$$

where $z_t(\psi)^T$ is a $(1 \times rp)$ vector, x_t a $(p \times 1)$ vector of explanatory variables, r is the number of regimes, and F_i is the transition function. We can express (A1) in matrix form:

$$\underset{(N\times 1)}{y} = \underset{(N\times rp)(rp\times 1)}{Z(\gamma, c)\beta} + \underset{(N\times 1)}{\varepsilon} \tag{A4}$$

where $\beta = (\beta_1^T \ldots \beta_r^T)^T$ and ε is *i.i.d.* $N(0, \sigma^2 I)$. The log likelihood function can be written:

$$\ell(\beta, \sigma^2, \psi) = constant - \frac{N}{2} ln(\sigma^2) - \frac{[y - Z(\psi)\beta]^T [y - Z(\psi)\beta]}{2\sigma^2}. \tag{A5}$$

In theory we can find the values of β, ψ and σ^2 which maximise the log likelihood function (A5). In practice, there are difficulties with the estimation procedure. The likelihood surface may be characterised by flat segments or numerous local maxima, both of which make the starting values of the parameters critical. The approach taken in this paper is to abbreviate or simplify the optimisation procedure by focussing on the ψ parameters, since they seem to be principal source of estimation difficulties. The optimisation problem is simplified by concentrating out the β and σ^2 parameters. The first order conditions with respect to β and σ^2 give the familiar OLS result:

$$\tilde{\beta}(\psi) = [Z(\psi)^T Z(\psi)]^{-1} Z(\psi)^T y \tag{A6}$$

and

$$\tilde{\sigma}^2(\psi) = \frac{[y - Z(\psi)\tilde{\beta}(\psi)]^T [y - Z(\psi)\tilde{\beta}(\psi)]}{N}. \tag{A7}$$

The optimisation procedure consists of maximising the concentrated log likelihood function,

$$\ell_c^* = -\frac{N}{2} ln(\tilde{\sigma}^2(\psi)), \tag{A8}$$

with respect to ψ. We can therefore estimate β and σ^2 by inserting the solution of (A8) into (A6) and (A7). The concentrated likelihood approach permits us to substantially simplify our optimisation problem by requiring that we limit our search, via a hill-climbing method, over the (ψ)-space. An additional advantage is that it reduces the parameter space sufficiently, permitting a grid-search approach in the (ψ,σ^2)-space. The starting values of ψ may be crucial, both in the hill-climbing method and in finding the global maximum given that the likelihood surface frequently has an erratic shape in RSMs.

We require, in the case of the STAR model, additional assumptions on the progression of the transition variable in order to forecast the value of the transition function. We assume in the case of the difference in the bank rate that the forecasted values are zero (no change in the bank rate from the last available observation). Our estimated STAR (bank) model is:

$$y_t = 3.805 + 0.402\, y_{t-1} - 0.340\, y_{t-2} + (-1.694 - 0.232\, y_{t-1} + 0.577\, y_{t-2}) \times F(\gamma, c) + \hat{\varepsilon}_t$$

$$\text{with } \hat{\gamma} = 175.434,\ \hat{c} = -0.330 \text{ and } \hat{\sigma} = 1.923.$$

Our STAR (US) model assumes that forecasted US GDP growth continues at the same rate that it averaged over the previous four quarters. Our estimated STAR (US) model is:

$$y_t = 4.648 + 0.194\, y_{t-1} + 0.010\, y_{t-2} + (-5.430 - 0.268\, y_{t-1} + 1.011\, y_{t-2}) \times F(\gamma, c) + \hat{\varepsilon}_t,$$

$$\text{with } \hat{\gamma} = 165.500,\ \hat{c} = 3.663 \text{ and } \hat{\sigma} = 1.902.$$

Hamilton Model. Hamilton (1989) proposed a RSM based on an unobservable state of nature S_t:[25]

[25] This two regime model can be expressed in the form of (5) by assuming: $F(S_t; \psi) = S_t$.

$$y_t = \alpha_0 + \alpha_1 S_t + z_t \tag{A9}$$

where

$$S_t = \begin{cases} 0 \text{ if regime } 0 \\ 1 \text{ otherwise} \end{cases}. \tag{A10}$$

The variable z_t in turn follows a AR(p) process which includes an error term:

$$z_t = \phi_1 z_{t-1} + \ldots + \phi_p z_{t-p} + \varepsilon_t. \tag{A11}$$

The interesting and innovative feature of the Hamilton approach is that the data generating mechanism is a function of the current, but unobservable state of nature, S_t. It is assumed that S_t follows a first-order Markov process. The first-order Markov assumption enables us to solve the maximisation of the joint likelihood function of the economic series and the unobserved states of nature. Though the states are unobservable, we can nonetheless infer the probability of being in one state by looking at previous observations.

The Modified Hamilton model refers to a standard Hamilton RSM in equation (A9), except the Modified version allows all of the AR coefficients to vary according to the states of nature:

$$z_t = \phi_{1S_t} z_{t-1} + \phi_{2S_t} z_{t-2} + \ldots + \phi_{pS_t} z_{t-p} + \varepsilon_t \tag{A12}$$

where $\phi_{iS_t} = \phi_{i0} + \phi_i S_t$ for i=1, 2, ..., p, and S_t is defined in (A10). This is similar to our STAR model, where the AR coefficients are function of the prevailing regime. We estimate this model by substituting (A9) into (A12), and then rewriting z_t in terms of lagged y_t and lagged states of nature:

$$y_t = \alpha_0 + \alpha_1 S_t + \phi_{1S_t}(y_{t-1} - \alpha_0 - \alpha_1 S_{t-1}) + \ldots + \phi_{pS_t}(y_{t-p} - \alpha_0 - \alpha_1 S_{t-p}) + \varepsilon_t \tag{A13}$$

Note that we can obtain, as a special sub-case, the Hansen (1992) model:

$$y_t = \alpha_0 + \alpha_1 S_t + \phi_{1S_t} y_{t-1} + \ldots + \phi_{pS_t} y_{t-p} + \varepsilon_t \tag{A14}$$

$\beta_1 = (\alpha_0 + \alpha_1, 1)$, $\beta_2 = (\alpha_0, 1)$ and $x_t = (1, z_t - \varepsilon_t)$.

where ϕ_{iS_t} and S_t are defined as before.

The estimation aspects of Hamilton's changing intercept model can be found in Hamilton (1989). The Modified Hamilton model is simply an extension of Hamilton's original changing intercept model and can be estimated in a similar manner.

The estimated Hamilton model is:

$$y_t = 3.064 + 1.779S_t + (0.457 - 0.524S_t)(y_{t-1} - 3.064 - 1.779S_{t-1})$$
$$+ (0.102 - 0.422S_t)(y_{t-2} - 3.064 - 1.779S_{t-2}) + (0.365 - 0.502S_t)(y_{t-3} - 3.064 - 1.779S_{t-3})$$
$$+ (-0.403 + 0.986S_t)(y_{t-4} - 3.064 - 1.779S_{t-4}) + \hat{\varepsilon}_t$$

$P(S_t = 1 \mid S_{t-1} = 1) = 0.339$, $P(S_t = 0 \mid S_{t-1} = 0) = 0.689$, $P(S_0 = 1) = 0.320$ and $\hat{\sigma} = 1.700$.

Note that $P(S_0 = 1) = \dfrac{1 - P_{00}}{2 - P_{11} - P_{00}}$ is the equilibrium probability of being in regime one.

Annex 2: Tables and Figures

Table A1: Root mean squared errors over various sample periods

1960-1995	Linear	STAR (bank)	STAR (US)	BT	Hamilton	HP Filter
Final						
One year	2.535	2.436	2.195	2.414	2.524	2.083
Two years	1.831	1.833	1.617	1.644	1.813	1.462
Three years	1.582	1.596	1.441	1.325	1.567	1.296
Four years	1.429	1.458	1.283	1.103	1.410	1.229
Five years	1.341	1.374	1.168	0.974	1.318	1.241
Quasi-Real						
One year	3.249	3.285	3.143	3.304	3.375	2.720
Two years	2.569	2.673	2.513	2.603	2.673	2.213
Three years	2.150	2.416	2.145	2.190	2.315	2.022
Four years	2.044	2.393	2.093	2.052	2.270	1.899
Five years	2.009	2.375	2.074	2.005	2.258	1.820
1965-1995						
Final						
One year	2.432	2.331	2.053	2.248	2.396	2.139
Two years	1.889	1.874	1.600	1.593	1.866	1.502
Three years	1.653	1.649	1.437	1.288	1.638	1.331
Four years	1.482	1.495	1.280	1.050	1.460	1.258
Five years	1.399	1.416	1.194	0.939	1.370	1.267
Quasi-Real						
One year	2.570	2.599	2.489	2.684	2.674	2.779
Two years	2.025	2.185	1.931	2.108	2.188	2.249
Three years	1.814	2.112	1.727	1.893	1.981	2.051
Four years	1.690	2.050	1.597	1.727	1.886	1.917
Five years	1.646	2.003	1.531	1.666	1.846	1.835
1970-1995						
Final						
One year	2.428	2.311	2.116	2.290	2.337	2.142
Two years	1.876	1.847	1.658	1.658	1.824	1.541
Three years	1.642	1.624	1.498	1.352	1.599	1.423
Four years	1.445	1.444	1.331	1.092	1.391	1.325
Five years	1.350	1.350	1.237	0.976	1.284	1.341
Quasi-Real						
One year	2.675	2.713	2.614	2.739	2.704	2.836
Two years	2.132	2.238	1.959	2.202	2.134	2.355
Three years	1.923	2.119	1.674	1.959	1.878	2.179
Four years	1.795	2.019	1.476	1.793	1.711	2.025
Five years	1.748	1.974	1.412	1.743	1.645	1.951

Figure A1: Quasi-real evolution of model fit (standard errors)

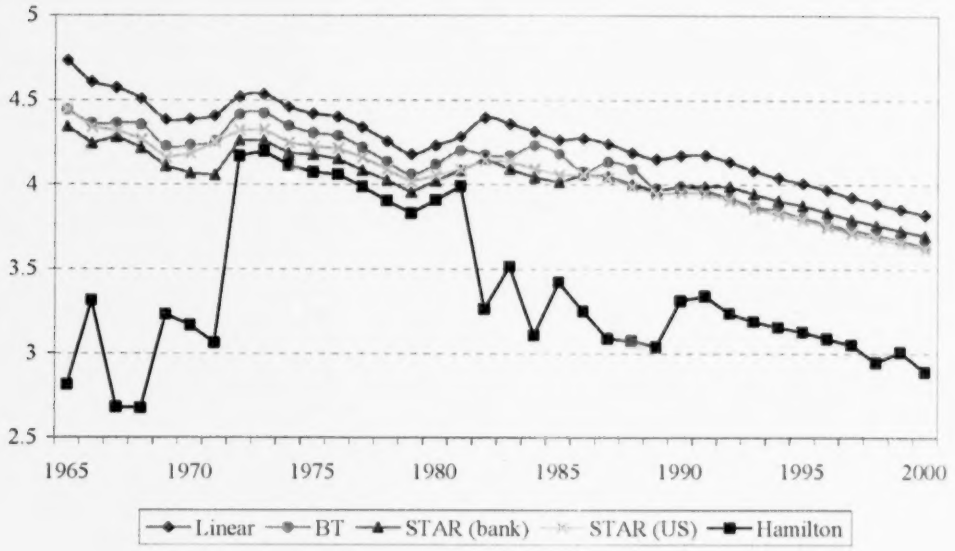


Figure A2: Quasi-real evolution of the ratio of the residual variance

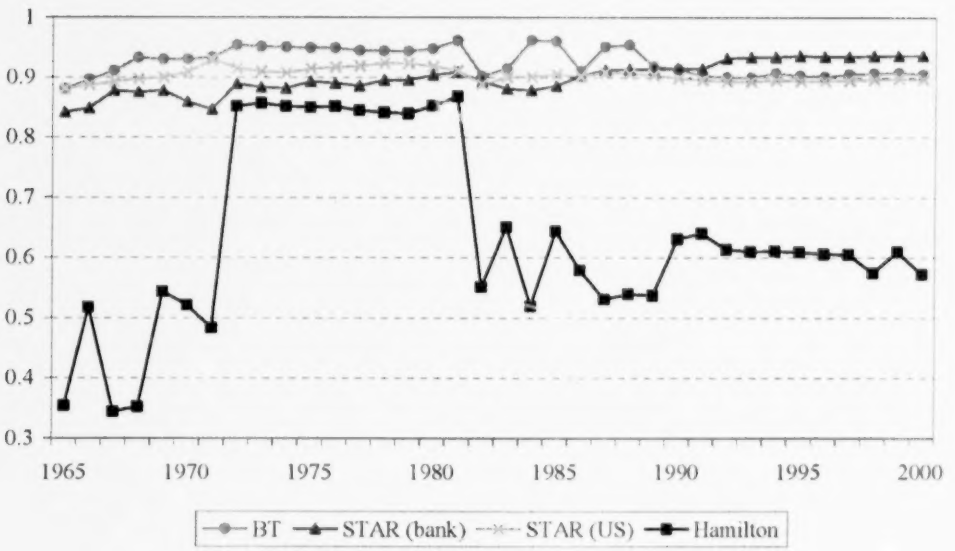


Table A2: Correlations between final rolling RMSEs

	Linear	BT	STAR (bank)	STAR (US)	Hamilton	HP
Linear	1	0.51	0.99	0.81	0.99	0.61
BT	0.51	1	0.42	0.72	0.42	0.80
STAR (bank)	0.99	0.42	1	0.74	0.99	0.54
STAR (US)	0.81	0.72	0.74	1	0.75	0.71
Hamilton	0.99	0.42	0.99	0.75	1	0.50
HP	0.61	0.80	0.54	0.71	0.50	1

Table A3: Correlations between quasi-real rolling RMSEs

	Linear	BT	STAR (bank)	STAR (US)	Hamilton	HP
Linear	1	0.92	0.98	0.95	0.97	0.72
BT	0.92	1	0.89	0.93	0.90	0.77
STAR (bank)	0.98	0.89	1	0.93	0.95	0.62
STAR (US)	0.95	0.93	0.93	1	0.94	0.76
Hamilton	0.97	0.90	0.95	0.94	1	0.78
HP	0.72	0.77	0.62	0.76	0.78	1